DOCUMENT RESUME

ED 186 453 TM 800 141

AUTHOR Sternberg, Robert J.
TITLE The Construct Validity of Aptitude Tests: An Information-Processing Assessment. Technical Report No. 20, (July 1, 1979 through September 30, 1979).
INSTITUTION Yale Univ., New Haven, Conn. Dept. of Psychology.
SPONS AGENCY Office of Naval Research, Washington, D.C. Personnel and Training Branch.
PUB DATE Oct 79
CONTRACT N0001478C0025
NOTE 39p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Abstract Reasoning; *Academic Aptitude; *Aptitude Tests; Aptitude Treatment Interaction; Cognitive Ability; *Cognitive Processes; Intelligence; *Learning Theories; *Task Analysis; *Test Validity
IDENTIFIERS Cattell Horn Fluid and Crystallized Ability Theory; *Construct Validity; Guilfords Structure of Intellect; Spearmans Two Factor Theory; Thurstones Primary Mental Abilities Schema

ABSTRACT

An information-processing assessment of the construct validity of aptitude tests is presented. It is stated that aptitude tests have been rather successful because they possess high construct validity, although the constructs in terms of which the tests may be most profitably understood are information-processing rather than psychometric ones. It is suggested that certain inconsistencies in the constructs proposed by alternative psychometric theories of aptitude tend to disappear when these constructs are understood in terms of more basic information-processing constructs. Several theories are discussed: componential theory; Spearman's two-factor theory; Thurstone's primary mental abilities; Cattell and Horn's fluid and crystallized ability; and Guilford's structure of intellect. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED186453

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Construct Validity of Aptitude Tests:

An Information-Processing Assessment

Robert J. Sternberg

Department of Psychology

Yale University

New Haven, Connecticut 06520

Technical Report No. 20
October, 1979

Approved for public release; distribution unlimited.
Reproduction in whole or in part is permitted for
any purpose of the United States Government.

This research was sponsored by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N0001478C0025, Contract Authority Identification Number NR 150-412.

IM800 141

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. 20	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) "The Construct Validity of Aptitude Tests: An Information-Processing Assessment"		5. TYPE OF REPORT & PERIOD COVERED Periodic Technical Report (1 Jul 79 - 30 Sep 79) 6. PERFORMING ORG. REPORT NUMBER Research Report No. 6-79
7. AUTHOR(s) Robert J. Sternberg		8. CONTRACT OR GRANT NUMBER(s) N0001478C0025
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Yale University New Haven, Connecticut 06520		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N; RR 042-04; RR 042-04-01; NR 150-412
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, Virginia 22217		12. REPORT DATE 1 Oct 79 13. NUMBER OF PAGES 33
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

To be published in the proceedings of a Colloquium on Construct Validation, Educational Testing Service, Princeton, October, 1979

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Information processing, construct validity, component, aptitude, factor

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This article presents an information-processing assessment of the construct validity of aptitude tests. It proposes that aptitude tests have been rather successful because they do possess high construct validity, although the constructs in terms of which the tests may be most profitably understood are information-processing rather than psychometric ones. Certain inconsistencies in the constructs proposed by alternative psychometric theories of aptitudes are proposed to disappear when these constructs are understood in terms of

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

more basic information-processing constructs.

The Construct Validity of Aptitude Tests:

An Information-Processing Assessment

Robert J. Sternberg

Yale University

Running head: Information-Processing Assessment

Send proofs to Robert J. Sternberg
Department of Psychology
Yale University
Box 11A Yale Station
New Haven, Connecticut 06520

Abstract

This article presents an information-processing assessment of the construct validity of aptitude tests. It proposes that aptitude tests have been rather successful because they do possess high construct validity, although the constructs in terms of which the tests may be most profitably understood are information-processing rather than psychometric ones. Certain inconsistencies in the constructs proposed by alternative psychometric theories of aptitude are proposed to disappear when these constructs are understood in terms of more basic information-processing constructs.

The Construct Validity of Aptitude Tests: An Information-Processing Assessment

The closing months of 1979 provide an excellent time for yet another assessment of the construct validity of the available range of aptitude tests, because during this decade the construct validity of aptitude tests has undergone such intense scrutiny. These months provide a particularly appropriate time for an information-processing assessment of construct validity, because the information-processing view of construct validation has become so salient in the seventies. Experimental psychologists finally took to heart Cronbach's (1957) plea for a unification of the two disciplines of scientific psychology, and let come to pass Thurstone's (1947) prediction that "the rough factorial map of a new domain will enable us to proceed beyond the exploratory factorial stage to the more direct forms of psychological experimentation in the laboratory" (p. 56).

I will divide my assessment of the construct validity of aptitude tests into four parts. In the first, I will motivate the assessment by discussing why, to some at least, an information-processing assessment of the construct validity of aptitude tests has seemed like such a good idea. In the second part, I will present a brief historical overview of some of what has been done toward this goal in the last decade, and will assess some successes and failures of the attempts that have been made. A more detailed overview is provided by Carroll and Maxwell (1979). In the third part, I will briefly describe my own present componential view

of aptitude. In the last part, I will explain what this view tells us about the construct validity of aptitude tests.

Psychometric Approaches to Construct Validation: Something's Missing

Psychometric approaches to construct validation of aptitude tests have generally sought to understand tests and test performance in terms of some kind of structural model, such as the factor-analytic model or the latent-trait model. Individual-differences variation in test performance is explained as deriving from one or more underlying structures, such as a general factor and specific factors, or a set of primary factors. These models have been criticized, rightly or wrongly, for a variety of sins, some of them venial, others, mortal. But for the present purposes, two aspects of the models are critical.

First, there is an element of circularity to these explanations that can become vicious when the explanations are used in construct validation. The aptitude tests under scrutiny were constructed on the basis of explicit or implicit psychometric theories of aptitude. These tests are supposed to provide construct validation of the aptitude theories, while at the same time the aptitude theories are supposed to provide construct validation of the aptitude tests. It's hard to have it both ways, if one's goal is an independent appraisal of the tests, of the theories, or, as is the case in construct validation, of both. The tests and theories may be conceptually compatible with each other, but neither can provide independent support for the other, because the interdependence between the two was established at the time the theories and tests were

created. What would seem to be needed for construct validation of both the tests and the theories is some different kind of aptitude theory that is logically (although one would hope, not psychologically) independent of the psychometric tradition of theories and tests. Information-processing theories and tests seem to fill this bill.

Second, aptitude tests require and aptitude theories deal with dynamic performance: Solution of test items occurs over time, and the construct validity of both tests and theories will depend in part upon whether they can account for what happens over time. But although aptitude tests requires dynamic performance, the indices of performance they provide are essentially static measures of terminal behavior on test items; and although aptitude theories must explain dynamic constructs, they generally explain these constructs in a static way, through structural models such as factor analysis. This does not mean that the theories are wrong or the tests inadequate. It does mean, though, that a complementary kind of theory and index of test performance are needed to deal with dynamic performance. The information-processing approach seems to provide what is needed. So let us turn to a consideration of just what we mean by information-processing theories and test indices.

Information-Processing Approaches to Construct Validation: An Historical Overview

What is an information-processing approach to construct validity? Indeed, what is an information-processing analysis? An information-processing analysis is an attempt to account for individuals' performance on one or more tasks in terms of the elementary processes used in

task performance, the strategies for task performance into which these processes combine, and the representations for information upon which the processes and strategies act. An information-processing approach to construct validation seeks to examine the validity of psychometric tests and theories of aptitude through an analysis of the information processing that occurs during test performance or that gives rise to a source of individual differences such as a factor. From an information-processing point of view, then, factors can be understood in information-processing terms (Carroll, 1976; Sternberg, 1977): The two viewpoints are not inconsistent, but complementary.

Information-processing analysis of aptitude test performance during the seventies has arisen in part as a reaction to psychometric analysis of test performance (see Hunt, Frost, & Lunneborg, 1973; Sternberg, 1977). But if one goes back a decade further to the origins of information-processing analysis, one would find that it originated in large part as a reaction to stimulus-response analysis of task performance (see Miller, Galanter, & Pribram, 1960). At the time, many experimental psychologists had become dissatisfied with the inability of stimulus-response theories to account for complex forms of behavior. These theories seemed to account very well for the contingencies that brought about behavior, and for the responses that eventuated from these stimuli; but they were less well able to account for the processes that took place in the head and that served to link external stimuli and responses. Psychometric and stimulus-response accounts of complex behavior differ in many, many respects, but they have

in common their seeming inability to provide dynamic accounts of information processing. As its name implies, this is where the information-processing approach excels. What forms have information-processing analyses of aptitudes taken, and how have they functioned in the construct validation of aptitude tests?

The Aptitude-Treatment Interaction Approach

The aptitude-treatment interaction (ATI) approach dates back past the beginning of the seventies, at least to Cronbach's (1957) APA presidential address, where the ATI approach was suggested as one means of reuniting the two disciplines of scientific psychology. The approach provides construct validation by assuming that if an aptitude is psychologically real in some sense, and if a given test actually measures that aptitude, then the measured aptitude should interact with at least one instructional treatment in a meaningful way. For example, it might be proposed that high-reasoning subjects would do better to learn by discovery, whereas low-reasoning subjects would do better to learn by rules. If a given aptitude does not show a meaningful pattern of interaction with any instructional treatment, then this lack of empirical consequences for that aptitude (as measured) might lead one to question why the aptitude should be labeled an "aptitude" at all. Similarly, if two aptitudes never interact differentially with various instructional treatments, then the lack of differential empirical consequences for the two aptitudes might lead one to question whether they should be identified as separate aptitudes.

The history of ATI research has been exhaustively documented by

Cronbach and Snow (1977): Hundreds of studies have failed to yield highly replicable findings, or to shed much light on the nature of aptitudes. Cronbach and Snow have suggested a number of technical inadequacies in past studies that are sufficient to account for the failures of so many ATI studies. But from our present vantage point, two interrelated elements seem to be key.

First, it has never been entirely clear how aptitude-treatment interaction research could be expected to yield a theory of aptitudes that would provide an independent basis for the construct validation of existing tests and theories. Rather, ATI methodology seems to provide a device for testing extant theories. It does not provide a method for decomposing test performance into elementary constituents of some kind, but rather a method for understanding overall test performance in an experimental, or, ideally, classroom context. Methodologies are needed for both functions, so this is certainly not a criticism of ATI methodology. It is merely a suggestion that ATI methodology does not provide the new, independent kinds of constructs needed for construct validation of psychometric tests and theories. Ideally, these new kinds of constructs could be employed in an ATI setting.

Second, if the pre-existing constructs that are used in ATI research are not useful for the purpose to which they are being put, then the research is likely to fail. Aptitude-treatment interaction experiments generally involve some instructional (or other experimental) treatment administered over time. The aptitude constructs used in the experiments should therefore be dynamic ones that are sensitive to the processes as well as

to the products of performance. But the constructs used in ATI research have been, for the most part, psychometric ones that are not very sensitive to dynamic changes in information processing over the duration of the treatment. These psychometric constructs, no matter how useful they may be for other purposes, are probably not well suited for use in these experiments. A factorially-based measure of spatial ability, for example, will not be optimal if performance in response to a treatment involves spatial processes at some points in time, but, say, linguistic processes at other points in time (as is the case with certain types of syllogisms). If this argument is correct, ATI research should be more successful if dynamic, information-processing constructs are used instead of psychometric ones, assuming, of course, that the particular information-processing constructs used are theoretically appropriate and measured adequately. Two recent studies along these lines suggest that information-processing constructs can result in clearcut aptitude-strategy interactions (Sternberg & Weil, in press; Mathews, Hunt, & MacLeod, Note 1).

The Cognitive-Correlates Approach

What Pellegrino and Glaser (1979) have referred to as the "cognitive-correlates approach" can be traced back to an article by Hunt, Frost, and Lunneborg (1973). This line of research has been continued by Hunt and his associates throughout the course of the seventies (see Hunt, 1978; Hunt, Lunneborg, & Lewis, 1975). The basic idea is that performance on complex aptitude tests can be understood in terms of the information-processing components that are used in the performance of relatively simple laboratory

tasks studied by cognitive psychologists. For example, one such task requires subjects to compare whether two letters are both in the same case (upper-upper or lower-lower) or each in a different case (upper-lower or lower-upper). The subject would be timed, say, for how long it takes him or her to determine that "a A" represent different cases (cf. Posner & Mitchell, 1967). Another such task requires subjects to indicate whether a given target letter (or number) occurs in a short string of letters (or numbers). The target may be presented either before or after the string. The subject might be asked, say, to memorize the letter string A-N-C-U-E." The subject would then be presented with a target letter, such as "U," and be asked to say as quickly as possible whether the letter "U" appears in the memory set. In both this task and the preceding one, time to respond is the major dependent variable. Using single tasks such as these to investigate verbal aptitude, Hunt et al. (1975) concluded that verbal aptitude can be understood largely in terms of the speed of access and retrieval of verbal information from short-term memory.

This kind of research has the potential for providing the kinds of constructs that are needed for construct validation of aptitude tests and theories. The cognitive theories under investigation are certainly independent of the psychometric tradition. The tasks were also developed independently of this tradition, however, and one might wonder whether they are the right tasks, and the processes that enter into them, the right processes. I find it very difficult to believe, on merely intuitive grounds, that the simple tasks used in the cognitive psychologist's

laboratory bear much relation to the complex tasks found on aptitude tests, or to the still more complex tasks required in everyday and in academic life. Although the laboratory cognitive tasks are quite nontrivial when examined in the context of certain cognitive theories, they appear to be rather trivial when viewed as a basis of performance for, say, comprehension of complex reading material. The experimental data Hunt and others have collected can be interpreted as supporting my reservations. Factor analyses of the aptitude scores and measures from the laboratory tasks reveal a major group factor comprising the various cognitive tasks (see Hunt et al., 1975). An interbattery factor analysis (Tucker, 1958) might be more illuminating, in that it would highlight only sources of variation shared across modalities (i.e., tasks versus tests). But I doubt the basic facts of the matter would change, since the aptitude tests tend to be rather highly intercorrelated, whereas the cognitive tasks are not very highly correlated with the aptitude tests, nor, even, for the most part, with each other.

The Task-Analysis Approach

The task-analysis approach in experimental psychology dates back to Donders (1868); in the study of aptitudes, people date it back in different ways, but I see it as having originated with Johnson's (1960) method of serial analysis. In task analysis, the task of interest is decomposed into smaller constituents of information processing, usually, component processes. The basic idea is that the way to understand complex aptitudes is to take the test items that are used to measure these aptitudes, and to decompose them into their basic component processes. Different methods have been proposed for decomposing tasks, including my own method of "componential analysis" (Sternberg, 1977, 1978a, 1979a). Componential analysis was self-

consciously created with construct validation in mind: "From a differential viewpoint, componential analysis may be viewed as a detailed algorithm for construct validation" (Sternberg, 1977, p. 65). I will not discuss each of the steps of the algorithm here, since they have been presented in some detail elsewhere (Sternberg, 1977). The major steps involve decomposing a complex aptitude item into a set of component processes, and relating these processes to global scores on reference ability test.

Consider, for example, an aptitude item such as an analogy, and the construct of "analogical reasoning" that this item is supposed to measure. Factorial analysis would seek to understand performance on analogy items by relating this performance to performance on other kinds of test items (say, series completions and classifications), and by finding an underlying factor that serves as a latent source of individual differences on all of these items. Cognitive-correlates analysis would seek to understand performance on analogy items by relating it to performance on very simple laboratory tasks requiring manipulation of information in short-term memory, attention to perceptual properties, and so on. Task analysis would seek to understand it by decomposing analogy performance into a set of underlying component processes. I have suggested, for example, that an analogy such as "LAWYER is to DOCTOR as CLIENT is to (A. PATIENT, B. MEDICINE)" is solved by (a) encoding each analogy term, retrieving from long-term memory a set of semantic attributes relevant to each individual term; (b) inferring the relation between LAWYER and DOCTOR; (c) mapping the higher-order relation from the first half of the analogy (headed by LAWYER) to the second half of the analogy (headed by CLIENT);

(d) applying from CLIENT to each of the two answer options, PATIENT and MEDICINE, the relation that was inferred from LAWYER to DOCTOR: (e) justifying one response as preferable to the other, even if nonideal; and (f) responding. These same components have been shown to be involved in the solution of other kinds of induction problems as well (Sternberg & Gardner, Note 2).

Task analysis, like cognitive-correlates analysis, provides an independent kind of theoretical construct for the construct validation of psychometric tests and theories. How well does the analysis fare? The question must be answered in two parts. First, we need to consider the internal validity of the analysis: How well do the proposed models of information processing in the tasks fit data collected from those tasks? Second, we need to consider the external validity of the analysis: How highly do task parameters correlate with external measures such as aptitude test scores, and are the patterns of correlation sensible?

From the standpoint of internal validation, the results of task analyses conducted during the past several years have been quite impressive. Models of task performance, although not accounting for all of the true variance in the response-time or response-choice data, have accounted for high proportions of the true variance in these data (e.g., Mulholland, Pellegrino, & Glaser, in press; Royer, 1971; Sternberg, 1977, in press-b). From the standpoint of external validation, however, the results are more mixed. Major performance components of information processing--such as the inference, mapping, and application components mentioned previously--

have shown generally significant correlations with aptitude test scores (when they are applied to nontrivial analogies or other induction items), but the magnitudes have not been impressive. The response component, which one would not expect to be particularly highly correlated with aptitude test scores, has shown correlations that are both significant and of high magnitude. Moreover, this result has proven to be replicable in a number of different studies of different tasks by different investigators. It seems likely that it is not response itself, but something confounded in the response component, that is responsible for the high correlations. Response is estimated as a regression constant, meaning that the response parameter includes in its duration any operation whose duration is constant across all of the items being solved. Apparently, some of the most interesting sources of individual-differences variation have been lumped into a "constant."

Thus, although task analyses of the kinds that have been performed seem to be a reasonable start toward the construct validation of aptitude tests and theories, they are far from the finish. The realization that the task analyses I and others were performing just weren't extracting the most interesting sources of individual-differences variation led me to propose an augmented form of my own brand of task analysis, componential analysis, that extracts kinds of components of information processing that were not extracted in the earlier analyses. Some of the new methods my collaborators and I have proposed for extracting the various kinds of components are described in Sternberg (1978b) and in Sternberg (1979b). In the next

section, I will briefly describe the theory underlying these new methods. This theory is proposed to apply to individuals across the whole range of levels of intelligence (Sternberg, 1979a, in press-a).

The Componential Theory of Intellectual Aptitudes: A Brief Overview

The proposed theory of intellectual aptitudes is component-based. Kinds of components can be classified in two different ways: by level of generality and by function (Sternberg, Note 3).

Levels of Generality

Components can be classified in terms of three levels of generality: General components are required for performance of all tasks within a given task universe; class components are required for performance of a proper subset of tasks (including at least two tasks) within a task universe; and specific components are required for the performance of single tasks within the task universe. I will ignore specific components, which seem not to carry their weight in a theory of intellectual abilities. Consider, for example, the components of analogical reasoning mentioned earlier. Suppose the universe of tasks is reasoning tasks of the kind found on many standardized tests of intellectual aptitudes. Encoding and response would seem to be general components, in that it is always necessary to encode the item information when it is presented, and to offer some kind of response after information processing on the item has been completed. Inference, mapping, application, and justification would seem to be class components. They are common to most inductive-reasoning tasks, i.e., tasks requiring individuals

to make predictions of some kind about new information on the basis of old information, but are not found on most deductive-reasoning tasks, i.e., tasks requiring individuals to produce or select a single logically certain solution.

The three kinds of components may sound something like three kinds of factors: general, group, and specific factors. Components differ in at least three important ways from factors, however. First, their psychological reference is more clear than, and quite different from, that of factors. A component is a process with an experimentally determinable duration, probability of error, and effect upon response choice. The psychological reference of a factor, on the other hand, has never been very clear (see Sternberg, 1977). Just what does it mean for a factor, or anything else, to be a "latent source of individual differences"? Components are believed to be latent sources of individual differences, but their psychological properties can nevertheless be clearly specified. Second, components are identified in a very different way from factors. Components for an individual subject are identified by decomposing response-time, error, or response-choice data for that subject on a particular task. The outcome is independent of whatever other subjects or tasks happen to be used in a given experiment. The components contribute (additively or otherwise) to solution time, error rate, and response choice for a given type of task. Factors, on the other hand, are usually identified on the basis of individual differences in the data of multiple subjects. The outcome is dependent upon both who is in the sample and what particular tasks are being factor analyzed. Third, compo-

nential solutions are unique, whereas factorial solutions are usually subject to arbitrary rotation (unless they were identified through confirmatory maximum likelihood methods).

Function

Components perform (at least) five different kinds of functions. Each of these five functions can be crossed with the three levels of generality, yielding 15 different types (overall) of components. Five of these types of components--all specific ones--are not of much interest, leaving ten interesting kinds of components in the theory.

Metacomponents are higher-order control processes that are used for planning how a problem should be solved, for making decisions regarding alternative courses of action during problem solving, and for monitoring solution processes. These metacomponents are sometimes referred to collectively as an "executive" or a "homunculus." Consider two examples of metacomponents, and of how they might function in an analogical-reasoning task. A first example is the selection of lower-order components: In the analogies task, an individual must select a set of component processes (such as the set presented earlier--encoding, inference, mapping, application, justification, response) that is sufficient for solution of the task. A second example is generation or selection of a strategy for combining these components: In the analogies task, an individual must decide how to sequence the set of components--for instance, should inference precede mapping, follow it, or perhaps occur simultaneously with it? Other metacomponents include selection of a representation for information,

selection of a speed-accuracy tradeoff, and monitoring of one's solution processes.

Performance components are used in the execution of various strategies for task performance. The component processes named above for the analogies task are examples of performance components. Performance components may be viewed as actualizing the plans laid down by the metacomponents. For example, once the decision has been made that the relation between the first two terms of an analogy must be inferred, the inference component actually performs the job of inferring this relation.

Acquisition components are skills involved in learning new information; retention components are skills involved in retrieving information that has been previously acquired; transfer components are skills involved in carrying over retained information from one situational context to another. New information is always presented in some kind of a context, no matter how impoverished. In terms of situational contexts, acquisition components represent particular skills involved in utilizing context cues to learn new information; retention components represent particular skills involved in retrieving the information that was once learned; and transfer components represent particular skills involved in relating old contexts to new ones. Consider, for example, one aspect of multiple contexts, their variability across multiple presentations of a single piece of information, such as the appearance of an unfamiliar word. Some things that make an individual more "highly verbal," according to the present theory, are, first, his or her greater ability to use this variability of context to acquire the meaning of the word; second, the individual's greater ability

to retrieve one or more contexts when the individual needs to retrieve the meaning of the word; and, third, the individual's greater ability to carry over a context, which has now become an "old" context, to a new context to understand the meaning of the word when it is presented in an entirely new context.

Interrelations Among Kinds of Components

Kinds of components are interrelated in various ways. Consider first how components of different levels of generality are related, and then how components serving different functions are related.

General components are related across tasks by virtue of the fact that a general component occurs, by definition, in the performance of each task in the task domain under consideration. At the other extreme, specific components are each unique to individual tasks. Thus, most of what will cause some tasks to be more or less closely related to other tasks is the class components they do or do not share. Classes of tasks such as inductive-reasoning tasks, deductive-reasoning tasks, spatial visualization tasks, etc., tend to share many class components within the members of each class, but fewer class components across classes of tasks. Some classes of tasks will be more closely related than others by virtue of their sharing more class components with each other. (A more detailed description of just how the tasks are interrelated via class components can be found in Sternberg, 1979a, and in Sternberg, Note 3.)

The various functional kinds of components are related to each other through the metacomponents. In the proposed system, metacomponents can directly activate, provide information to, and receive feedback from each

other kind of component. A given metacomponent can also communicate with and activate other metacomponents. The other kinds of components can activate and communicate with each other only indirectly through the metacomponents, however. Thus, all control and information in the system pass directly to and from the metacomponents: The metacomponents act as a filter in the relations of the other kinds of components to each other (see Sternberg, Note 3). For example, acquisition of information affects retention of information, but only via the link of acquisition and retention components to metacomponents, not via any direct link of acquisition and retention components to each other.

The metacomponents are able to process only a limited amount of information at a given time. In a difficult task, and especially in a new and difficult one, the amount of information being fed back to the metacomponents may exceed their capacity to act upon this information. In this case, the metacomponents become overloaded, and valuable information that cannot be processed may simply be wasted. The total information-handling capacity of the metacomponents of a given system will thus be an important limiting aspect of the system. Similarly, capacity to allocate attentional resources so as to minimize the probability of bottlenecks will be part of what determines the effective capacity of the system (see also Hunt, Note 3).

Construct Validation of Aptitude Tests via the Componential Theory

The componential theory of intellectual aptitudes sketched in the previous section can provide a basis for the construct validation of aptitude tests and the factorial theories upon which these tests have been based. I propose that what seem like conflicts between theories disappear when these

theories (and the tests they have generated) are viewed through the componential lens.

How is it possible for more than one psychometric theory of intelligence to be correct? It is possible because these theories differ from each other primarily in terms of the way a given factorial solution is rotated. The choice of a rotation is mathematically arbitrary, although different rotations of a given factorial solution seem to highlight different psychological elements. I have claimed previously that the choice of a rotation is a matter of convenience (Sternberg, 1977). It depends upon what distribution of components one wishes to highlight. Consider how various psychometric theories can be interpreted in terms of the particular set of components they render salient.

Spearman's Two-Factor Theory

According to Spearman's (1927) two-factor theory and tests derived from it, intelligence can be understood in terms of two kinds of factors: a general factor whose influence permeates all tests of intellectual aptitude, and specific factors that are each found only on single aptitude tests or tasks. Evidence in support of this theory tends to be obtained in unrotated factorial solutions. According to the componential theory, individual differences in the general factor are attributable to individual differences in the effectiveness with which general components are performed. In other words, the general factor comprises a set of general components that is common to a wide variety of intellectual tasks. Specific factors comprise specific components.

As it happens, the metacomponents have a much higher proportion of general components among them than do any of the other kinds of components, presumably because the executive routines needed to plan, monitor, and possibly replan performance are highly overlapping across tasks of a widely differing nature. Thus, individual differences in metacomponential functioning will be primarily responsible for the appearance of individual differences of a general nature. Metacomponents are not solely responsible for the appearance of a general factor, however. Most behavior, and probably all of the behavior exhibited on intelligence tests, is learned. Thus, there are certain to be acquisition components, as well as retention and transfer components, whose past influence will have a present effect upon individual differences in general ability. Finally, certain performance components, such as encoding and response, may be general to a wide variety of tasks, and thus also have an effect upon individual differences in the general factor.

Thurstone's Theory of Primary Mental Abilities

According to Thurstone's (1938) theory of primary mental abilities and the tests derived from it, intelligence can be viewed as comprising a small set of multiple factors, or primary mental abilities: verbal comprehension, number, spatial visualization, word fluency, perceptual speed, reasoning, and memory. A multiple factor solution such as this one tends to appear when factorial solutions are rotated to "simple structure." Simple structure solutions, like unrotated solutions, seem to have a special appeal to psychometricians, and I believe there is a psychological

basis for this appeal. Whereas the unrotated solution seems to provide the best overall measure of individual differences in general components, a simple structure solution seems to provide the best overall measure of individual differences in class components: There seems to be minimal overlap in class components across factors when factors are rotated in this way. Although overlap among class components is minimized, one would nevertheless expect some correlation between pairs of primary mental abilities, due to their overlap in general components. Indeed, when simple-structure solutions are factored, they tend to yield a second-order general factor, which I believe captures the shared variation across primary mental abilities due to general components.

Consider, for example, the inductive-reasoning factor that appears in Thurstone's theory. I have claimed that a relatively small set of class performance components--inference, mapping, application, and justification--appear in a wide array of inductive reasoning tasks (Sternberg, 1979a). Studies of spatial ability suggest that a rather small set of class performance components can probably account for performance on a variety of spatial tasks (see, e.g., Shepard & Metzler, 1971). And our present research on verbal comprehension suggests the possibility that a relatively small set of class acquisition, retention, and transfer components may account for much of the individual-differences variation in verbal comprehension tasks.

Cattell and Horn's Theory of Fluid and Crystallized Ability

Cattell and Horn (see, e.g., Cattell, 1971; Horn, 1968) have proposed

that general ability can be divided into two subabilities, fluid ability and crystallized ability. Fluid ability is best measured by tests of abstract reasoning, such as visual analogies, classifications, and series problems. Crystallized ability is best measured by tests of vocabulary, reading comprehension, and general information. A very similar theory has been proposed by Vernon (1971), whose major group factors of practical-mechanical ability and verbal-educational ability seem to correspond closely, if not exactly, to Cattell and Horn's factors of fluid and crystallized ability. Factors such as these seem to result from hierarchical forms of factor analysis. On the componential view, crystallized ability tests seem best to separate the products of acquisition, retention, and transfer components, whereas fluid ability tests seem best to separate the execution of performance components. I believe that the measurement of crystallized ability in typical testing situations differs in a key way from the measurement of fluid ability. Whereas the measurement of crystallized ability involves accumulated products of past executions of components of acquisition, retention, and transfer, measurement of fluid ability involves current execution of components of performance.

The difference in the way the two sets of components are measured may explain why measured crystallized ability tends to increase indefinitely with increasing age (short of senility), whereas fluid ability starts to decline in early to middle adulthood. Products of performance are already established and unlikely to show deterioration except through the effects of senility. Current execution of performance components or

any other kinds of components is likely to deteriorate with age, however.

Guilford's Structure-of-Intellect Theory

Guilford's (1967) structure-of-intellect theory claims that intelligence can be viewed as comprising 120 distinct intellectual aptitudes, each represented by an independent factor. Support for this theory can be obtained by Procrustean rotation of factorial solutions toward the prespecified theory. Horn and Knapp (1973) have cast a shadow over the psychometric, and hence psychological, acceptability of this theory. Nevertheless, I believe that there probably is a psychological basis to at least some aspects of this theory, and that these aspects of the theory can be interpreted in componential terms.

A given component must act upon a particular form of representation for information, and upon a particular type of information. The representation, for example, might be spatial or linguistic; the type of information (content) might be an abstract geometric design, a picture, a symbol, a word, etc. Forms of representation and contents, like components, can serve as sources of individual differences: A given individual might be quite competent when applying a particular component to one kind of content, but not when applying it to another. Representation, content, and process have been more or less confounded in most factorial theories, probably because certain components tend more often to operate upon certain kinds of representations and contents, and other components tend more often to operate upon different kinds of representations and contents. This confounding serves a practical purpose of keeping to a manageable

number the factors appearing in a given theory or test. But it does obscure the probably partially separable effects of process, representation, and content. Guilford's theory self-consciously tries to provide some separation, at least between process and content. I doubt the product dimension has much validity, other than through the fact that different kinds of products probably involve slightly different mixes of components. On the one hand, the theory points out the potential separability of process and content. On the other hand, it does so at the expense of manageability. Moreover, it seems highly unlikely that the 120 factors are independent, as they will show overlaps, at minimum, in shared metacomponents.

Conclusions

To conclude, the componential theory can provide at least a tentative and sketchy account of how different forms of factor analysis and rotation can support different factorial theories of intelligence. On this view, each so-called theory can be viewed as a special case of a single theory. Each special case highlights different constellations of individual differences, and which constellation of individual differences is important will depend upon the purpose the theory or test is intended to serve. In an important sense, then, most of the theories are "correct." The componential theory provides a complementary perspective on the various differential theories.

Tests of intellectual aptitude have numerous failings, but when all is said and done, there are very few nontrivial predictions of any kind that we can make in life that are a great deal more accurate than those made by

aptitude tests. Many people, in fact, are willing to acknowledge that aptitude tests have been the most successful technological innovation that has arisen from psychological theory and research. If this is merely an unhappy commentary on what psychology can produce, so be it. But I think this is not the case. During the seventies, experimental psychologists have embarked upon an intensive effort to examine the construct validity of tests and theories of intellectual aptitude. My retrospective view of the outcomes of this research is that it has supported rather than refuted their construct validity. Indeed, we knew the tests were working rather well; now we have, perhaps, a better idea of why. This is not to say that the tests are perfect, although I suspect that many of the imperfections that people complain about most bitterly reside not in the tests, but in people's sometimes unrealistic notions about what the tests can do. There is much more to intelligence than is measured by the relatively narrow focus of current aptitude tests, and there is much more to academic and life achievement than could be measured by any kind of intelligence test. If the validity coefficients of aptitude tests have not risen appreciably in the past thirty or forty years, it may be because we have reached the limits of what measurement of a narrow set of intellectual aptitudes can predict.

It is quite possible that eventually we may have tests that measure some of the kinds of components I have discussed above. Such tests would serve an important theoretical purpose that is not being served by current tests. They would be measuring performance at a more basic level

than is possible through existing test batteries. The question of practical value will not really be answerable until the tests are in a usable form, although I suspect such tests would supplement rather than replace the descendants of our present kinds of tests. In the meantime, I believe we would do well to continue the basic information-processing research we have been doing, since it seems to be bringing us closer to an understanding of what intelligence is, and of what intelligence tests measure.

Reference Notes

1. Mathews N.M., Hunt, E., & MacLeod, C. Strategy choice and strategy training in sentence-picture verification. Manuscript submitted for publication, 1979.

2. Sternberg, R.J., & Gardner, M.K. Unities in inductive reasoning (NR 150-412 ONR Technical Report No. 18) New Haven: Department of Psychology, Yale University, 1979.

3. Sternberg, R.J. Components of human intelligence. (NR 150-412 ONR Technical Report No. 19) New Haven: Department of Psychology, Yale University, 1979.

4. Hunt, E.B. Intelligence as an information processing concept. (NR 154-398 ONR Technical Report) Seattle: Department of Psychology, University of Washington, 1979.

References

Carroll, J.B. Psychometric tests as cognitive tasks: A new "structure of intellect." In L.B. Resnick (Ed.), The nature of intelligence. Hillsdale, N.J.: Erlbaum, 1976.

Carroll, J.B., & Maxwell, S.E. Individual differences in cognitive abilities. Annual Review of Psychology, 1979, 30, 603-640.

Cattell, R.B. Abilities: Their structure, growth, and action. Boston: Houghton-Mifflin, 1971.

Cronbach, L.J. The two disciplines of scientific psychology. American Psychologist, 1957, 12. 671-684.

Cronbach, L.J., & Snow, R.E. Aptitudes and instructional methods. New York: Irvington, 1977.

Donders, F.C. Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool, 1868-1869, Tweede reeks, II, 92-120.

Guilford, J.P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Horn, J.L. Organization of abilities and the development of intelligence. Psychological Review, 1968, 75, 242-259.

Horn, J.L., & Knapp, J.R. On the subjective character of the empirical base of Guilford's structure-of-intellect model. Psychological Bulletin, 1973, 80, 33-43.

Hunt, E.B. Mechanics of verbal ability. Psychological Review, 1978, 85, 109-130.

Hunt, E.B., Frost, N., & Lunneborg, C.E. Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), The psychology of learning and motivation (Vol. 7). New York: Academic Press, 1973.

Hunt, E.B., Lunneborg, C.E., & Lewis, J. What does it mean to be high verbal? Cognitive Psychology, 1975, 7, 194-227.

Johnson, D.M. Serial analysis of thinking. In Annals of the New York Academy of Sciences (Vol. 91). New York: New York Academy of Sciences, 1960.

Miller, G.A., Galanter, E., & Pribram, K.H. Plans and the structure of behavior. New York: Holt, Rinehart, & Winston, 1960.

Mulholland, T., Pellegrino, J.W., & Glaser, R. Components of geometric analogy solution. Cognitive Psychology, in press.

Pellegrino, J.W., & Glaser, R. Cognitive correlates and components in the analysis of individual differences. In R.J. Sternberg & D.K. Detterman (Eds.), Human intelligence: Perspectives on its theory and measurement. Norwood, N.J.: Ablex, 1979.

Posner, M., & Mitchell, P. Chronometric analysis of classification. Psychological Review, 1967, 74, 392-409.

Royer, F.L. Information processing of visual figures in the digit symbol substitution task. Journal of Experimental Psychology, 1971, 87, 335-342.

Shepard, R.N., & Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

Spearman, C. The abilities of man. New York: Macmillan, 1927.

Sternberg, R.J. Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, N.J.: Erlbaum, 1977.

Sternberg, R.J. Componential investigations of human intelligence. In A. Lesgold, J. Pellegrino, S. Fokkema, & R. Glaser (Eds.). Cognitive psychology and instruction. New York: Plenum, 1978. (a)

Sternberg, R.J. Isolating the components of intelligence. Intelligence, 1978, 2, 117-128. (b)

Sternberg, R.J. The nature of mental abilities. American Psychologist, 1979, 34, 214-230. (a)

Sternberg, R.J. A review of "Six authors in search of a character:" A play about intelligence tests in the year 2000. In R.J. Sternberg & D.K. Detterman (Eds.), Human intelligence: Perspectives on its theory and measurement. Norwood, N.J.: Ablex, 1979. (b)

Sternberg, R.J. Cognitive-behavioral approaches to the training of intelligence in the retarded. Journal of Special Education, in press. (a)

Sternberg, R.J. Representation and process in linear syllogistic reasoning. Journal of Experimental Psychology: General, in press. (b)

Sternberg, R.J., & Weil, E.M. An aptitude-strategy interaction in linear syllogistic reasoning. Journal of Educational Psychology, in press.

Thurstone, L.L. Primary mental abilities. Chicago: University of Chicago Press, 1938.

Thurstone, L.L. Multiple factor analysis. Chicago: University of Chicago Press, 1947.

Tucker, L.R. An inter-battery method of factor analysis. Psychometrika, 1958, 23, 111-136.

Vernon, P.E. The structure of human abilities. London: Methuen, 1971.

Footnote

Preparation of this report was supported by Contract N0001478C0025 from the Office of Naval Research. I am grateful to Janet Powell and Sandra Scarr for comments on an earlier version of this manuscript, and to my graduate-student collaborators in research, who have helped shape my views on the construct validation of aptitude tests.

Technical Reports Presently in this Series

NR 150-412, ONR Contract N00014-78-C-0025

#1. Sternberg, R. J. Intelligence research at the interface between differential and cognitive psychology: Prospects and proposals. January, 1978.

#2. Sternberg, R. J. Isolating the components of intelligence. January, 1978.

#3. Sternberg, R. J., Guyote, M. J., & Turner, M. E. Deductive reasoning. January, 1978.

#4. Sternberg, R. J. Toward a unified componential theory of human reasoning. April, 1978.

#5. Guyote, M. J., & Sternberg, R. J. A transitive-chain theory of syllogistic reasoning. April, 1978.

#6. Sternberg, R. J., & Turner, M. E. Components of syllogistic reasoning. April, 1978.

#7. Sternberg, R. J., Tourangeau, R., & Nigro, G. Metaphor, induction, and social policy: The convergence of macroscopic and microscopic views. April, 1978.

#8. Sternberg, R. J. A proposed resolution of curious conflicts in the literature on linear syllogistic reasoning. June, 1978.

#9. Sternberg, R. J. The nature of mental abilities. June, 1978.

#10. Sternberg, R. J. Psychometrics, mathematical psychology, and cognition: Confessions of a closet psychometrician. June, 1978.

#11. Tourangeau, R., & Sternberg, R. J. Understanding and appreciating metaphors. June, 1978.

#12. Sternberg, R. J. Representation and process in transitive inference. October, 1978.

#13. Tourangeau, R., & Sternberg, R. J. Aptness in metaphor. October, 1978.

#14. Sternberg, R. J. Contrasting conceptions of intelligence and their educational implications. November, 1978.

#15. Sternberg, R. J., & Weil, E. M. An aptitude-strategy interaction in linear syllogistic reasoning. April, 1979.

#16. Sternberg, R. J. Intelligence tests in the year 2000: What forms will they take and what purposes will they serve? April, 1979.

#17. Sternberg, R. J. New views on IQs: A silent revolution of the 70s. April, 1979.

Technical Reports Presently in this Series

NR 150-412, ONR Contract N0001478C0025

#18. Sternberg, R. J., & Gardner, M. K. Unities in inductive reasoning. October, 1979.

#19. Sternberg, R. J. Components of human intelligence. October, 1979.

#20. Sternberg, R. J. The construct validity of aptitude tests: An information-processing assessment. October, 1979.

#21. Schustack, M. W., & Sternberg, R. J. Evaluation of evidence in causal inference. October, 1979.